

EXPÉRIENCES PROFESSIONNELLES

Data Scientist – MACIF – Issy val de seine Avr. 2026 – Présent

- Conception et implémentation de modèles de scoring et de segmentation client.
- Réalisation d'études sur des campagnes marketing : analyse des comportements clients, mesure d'impact et restitution des insights aux équipes métier.
- Validation statistique des modèles (performance, stabilité, robustesse) et mise en œuvre de processus de data quality.

Outils : Python (Pandas, Scikit-learn, PyTorch), SQL, NoSQL, Dataiku DSS, Power BI, Git, outils Cloud

Data Analyst / Data Scientist – Brink's France Montrouge Juin – Août 2025 (2 mois)

- Analyse statistique des données transactionnelles pour identifier les facteurs influençant le délestage des automates.
- Modélisation de séries temporelles avec Prophet (R) pour la prévision des dates de délestage en fonction des flux financiers.
- Conception et automatisation d'un tableau de bord de suivi hebdomadaire des performances commerciales.

Outils : R (Prophet, ggplot2, dplyr), Power BI, SQL

PROJET PERSONNEL

Application mobile iOS/web En cours

- Développement full-stack d'une application Flutter (iOS/web) avec architecture modulaire et gestion d'état.
- Création d'un chatbot IA via l'API OpenAI (GPT-4) : mise en œuvre d'une architecture RAG, indexation de documents, recherche sémantique vectorielle et prompt engineering avancé.
- Sécurisation des appels API et gestion des abonnements via Firebase Cloud Functions (différenciation des fonctionnalités selon le plan utilisateur).
- Exploration de Vertex AI (embeddings) et Firestore Vector Search pour l'amélioration des capacités de recherche sémantique.

Outils : Flutter, Firebase, OpenAI API (GPT-4), RAG, Vertex AI, Firestore Vector Search, Prompt Engineering

PROJETS UNIVERSITAIRES

Modélisation de survie – Cancer du sein – Université Paris-Saclay / cBioPortal 2025

- Analyse clinique et moléculaire sur +2 000 patientes : biomarqueurs ER/PR/HER2, sous-types PAM50.
- Construction et comparaison de modèles de survie (Cox linéaire/non linéaire, RandomForestSRC).

Outils : R (survival, randomForestSRC), Python

Classification RNA-seq – scMARK – Université Paris-Saclay / RAMP 2025

- Pipeline de classification sur données déséquilibrées : PCA, augmentation de données, stratégie en deux étapes.
- Implémentation et comparaison de modèles Random Forest et XGBoost.

Outils : Python (Scikit-learn, NumPy, Pandas)

FORMATION

Master 2 Data Science – Spécialité Assurance (en cours)
Université Paris-Saclay – Stats bayésiennes, deep learning, NLP, LLM, analyse de survie.

Master 1 Mathématiques et Interactions
Université Paris-Saclay – ML, séries temporelles, optimisation, maths financières.

Licence Mathématiques – Université d'Évry

COMPÉTENCES

Langages Python (Pandas, NumPy, Scikit-learn, PyTorch), R, SQL, NoSQL, C/C++, Git, LaTeX

IA & ML ML, Deep Learning, NLP, LLM (GPT-4, CamemBERT), RAG, Prompt Engineering, Analyse de survie, Séries temporelles

Outils Dataiku DSS, Power BI, Firebase, Flutter, Vertex AI, Firestore, Cloud

Langues Français (natif) | Anglais (B2)